JOURNAL OF MULTIVARIATE ANALYSIS 8, 453–467 (1978)

Some Optimization Problems with Applications to Canonical Correlations and Sphericity Tests

C. G. KHATRI

Gujarat University, Ahmedabad, India

Communicated by P. R. Krishnaiah

Optimization problems are connected with maximization of three functions, namely, geometric mean, arithmetic mean and harmonic mean of the eigenvalues of $(X'\Sigma X)^{-1}X'\Sigma Y(Y'\Sigma Y)^{-1}Y'\Sigma X$, where Σ is positive definite, X and Y are $p \times r$ and $p \times s$ matrices of ranks r and s ($\geqslant r$), respectively, and $X'Y = 0$. Some interpretations of these functions are given. It is shown that the maximum values of these functions are obtained at the same point given by $X = (h_1 + \epsilon_1 h_p, \ldots, h_r + \epsilon_r h_{p-r+1})$ and $Y = (h_1 - \epsilon_1 h_p, \ldots, h_r - \epsilon_r h_{p-r+1}, y_{r+1}, \ldots, y_s)$, where $h_1, \ldots, h_p$ are the eigenvectors of Σ corresponding to the eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$, $\epsilon_j = +1$ or -1 for $j = 1, 2, \ldots, r$ and $y_{r+1}, \ldots, y_s$ are linear functions of $h_{r+1}, \ldots, h_{p-r}$. These results are extended to intermediate stationary values. They are utilized in obtaining the inequalities for canonical correlations $\theta_1, \ldots, \theta_r$ and they are given by expressions (3.8)–(3.10). Further, some new union–intersection test procedures for testing the sphericity hypothesis are given through test statistics (3.11)–(3.13).

1. INTRODUCTION

Let x and y be two nonsingular random vector variables with r and s ($s \geqslant r$) elements, respectively. Suppose their second moments exist. Then, the covariance matrices $V(x)$ and $V(y)$ of x and y are positive definite (p.d.). We shall define a square root $\Sigma^{1/2}$ of Σ, a p.d. matrix, by $\Sigma = (\Sigma^{1/2})(\Sigma^{1/2})'$. A correlation matrix between x and y is defined by

$$R = \{V(x)\}^{-1/2}\,\mathrm{Cov}(x, y)[\{V(y)\}^{-1/2}]'. \tag{1.1}$$

A measure of dependence or correlationship between x and y can be taken as a

Received July 11, 1977; revised April 3, 1978.

AMS 1970 subject classification: Primary 15A42; Secondary 62H20, 62H15.

Key words and phrases: Correlationship, eigenvalues and eigenvectors, efficiency, canonical correlations, multiple correlation, union–intersection test procedures, sphericity, geometric, arithmetic and harmonic means, inequalities.

0047-259X/78/0083–0453$02.00/0

proper function of the nonzero eigenvalues of RR'. We shall propose the following three measures,

$$g = | RR' |, \qquad a = \mathrm{tr}(RR')/r, \qquad \text{and} \qquad h = r/\mathrm{tr}\{(RR')^{-1}\}. \tag{1.2}$$

We can observe that $0 \leqslant h \leqslant g^{1/r} \leqslant a \leqslant 1$ and if the rank of $\mathrm{Cov}(x, y)$ is less than r, then $g = h = 0$, while a may not be zero. Further, when $r = 1$, RR' becomes the square of the multiple correlation between x and y, and $a = g = h$. For $r \geqslant 1$,

$$g = \prod_{i=1}^{r} \rho_i^2, \qquad a = \sum_{i=1}^{r} \rho_i^2/r, \qquad \text{and} \qquad h = r\Big/\sum_{i=1}^{r} \rho_i^{-2}, \tag{1.3}$$

where ρ_i^2 $(i = 1, 2,..., r)$ are the square of the canonical correlations between x and y. Note that $g^{1/r}$, a, and h are respectively the geometric, arithmetic, and harmonic means of the squares of the canonical correlations, while the measure for the likelihood ratio principle is $v^{-1} = \prod_{i=1}^{r} (1 - \rho_i^2) = | I - RR' |$.

Let us connect these measures to the idea of efficiency. Let y be the most efficient (or a maximum likelihood) estimate for the parameter θ and x be any other estimate of θ. Then, the efficiency of the estimate x can be taken as any one of g, a, or h. For example, let the general linear model be

$$y = X\beta + u, \tag{1.4}$$

where $E(u) = 0$, $V(u) = \Sigma$, a p.d. matrix, and X is a $p \times r$ matrix of rank r. Then, the best efficient estimate of β is

$$\beta = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y, \tag{1.5}$$

while the usual least-squares estimate of β is

$$b = (X'X)^{-1}X'y. \tag{1.6}$$

Then, the correlation matrix between b and β is

$$R = \{(X'X)^{-1}(X'\Sigma X)(X'X)^{-1}\}^{-1/2}\{(X'\Sigma^{-1}X)^{-1/2}\}',$$

and this gives

$$\begin{aligned} g &= | X'X |^2/| X'\Sigma X | \, | X'\Sigma^{-1}X |, \\ a &= \mathrm{tr}\{(X'X)(X'\Sigma X)^{-1}(X'X)(X'\Sigma^{-1}X)^{-1}\}/r \end{aligned} \tag{1.7}$$

and

$$h = r/\mathrm{tr}\{(X'X)^{-1}(X'\Sigma X)(X'X)^{-1}(X'\Sigma^{-1}X)\}.$$

The criterion for efficiency as taken by Bloomfield and Watson (1975) or Knott (1975) is g but we can take a and h as the criteria for efficiency. They have established the minimum value of g over the variation of X, and it occurs at $X = (h_1 + \epsilon_1 h_p, ..., h_r + \epsilon_r h_{p-r+1})$, where $h_1, ..., h_p$ are the eigenvectors of Σ with respect to the eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$ and $\epsilon_j = +1$ or -1 for $j = 1, 2, ..., r$.

Let us consider a p-vector w and let us consider two linear functions $x = X'w$ and $y = Y'w$ where X and Y are respectively $p \times r$ and $p \times s$ matrices of ranks r and s ($\geqslant r$). Suppose, $V(w) = \Sigma$ is a p.d. matrix. Then, the three measures of (1.3) are given by

$$g = | X'\Sigma Y(Y'\Sigma Y)^{-1} Y'\Sigma X |/| X'\Sigma X |, \tag{1.8}$$

$$a = \operatorname{tr}\{(X'\Sigma X)^{-1} X'\Sigma Y(Y'\Sigma Y)^{-1} Y'\Sigma X\}/r, \tag{1.9}$$

and

$$h = r/\operatorname{tr}\{X'\Sigma X[X'\Sigma Y(Y'\Sigma Y)^{-1} Y'\Sigma X]^{-1}\}. \tag{1.10}$$

When $r = s = 1$, Eaton (1976) has given the maximum values of $g = a = h$ under the variations of X and Y such that $X'Y = 0$, and he established an inequality between the canonical correlation and the maximum value of g. In this paper, we try to maximize g, a, and h under the variations of X and Y subject to $X'Y = 0$, and they occur at the same value of X as mentioned at the end of the previous paragraph. These results are connected to canonical correlations and they are used to obtain union–intersection test procedures for testing the sphericity hypothesis. These sphericity test procedures are different from those established by Venables (1976). This may be due to the fact that he used the "likelihood ratio" measure v for optimization. For obtaining a relationship for the intermediate roots, the idea of Amir–Moez and Ali (1956) is used, which uses the infimum and the supremum over two different regions. These results are given in Sections 2 and 3.

2. Some Theorems on Optimization

The matrices in this paper will be over the real space. The identity matrix will be denoted by I_n, $n \times n$ being the order of the matrix, or simply by I if there is no confusion. The null matrix will be simply denoted by 0. A matrix B will be said to be positive definite (p.d.) (or positive semidefinite (p.s.d.) if B is symmetric and for every vector x, $x'Bx > 0$ (or $\geqslant 0$). We shall reserve Σ for a p.d. matrix and its eigenvalues will be denoted by $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$. $\Sigma^{1/2}$ will be denoted as a square root of Σ such that $\Sigma = (\Sigma^{1/2})(\Sigma^{1/2})'$. Let $\mathscr{M}(p, r)$ be a class of $p \times r$ matrices of rank r ($\leqslant p$) and $\mathscr{M}(p, r, s)$ a class of

$p \times r$ and $p \times s$ matrices X and Y of ranks r and s, respectively, such that $X'Y = 0$. The differential of a function $f(X)$ is denoted by $(df(X)) = \sum_{i,j}(\partial f(X)/\partial x_{ij})(dx_{ij}) = \operatorname{tr}(\partial f(X)/\partial X)'(dX)$, where $(\partial f(X)/\partial x_{ij}) = \partial f(X)/\partial X$. Further, we may observe the following properties of the differentials: $(d(AB)) = (dA)B + A(dB)$, $(d \ln | A |) = \operatorname{tr}\{A^{-1}(dA)\}$ and $(dA^{-1}) = -A^{-1}(dA)A^{-1}$.

LEMMA 1. *Let A and B be two* p.d. *matrices and let $b_1 > b_2 > \cdots > b_q > 0$ be distinct eigenvalues of AB. Let $f(x)$ be a function of x such that* (i) *$x^t f(x)$ is a polynomial in x for some non-negative integer t and* (ii) *$f(b_i) \neq f(b_j)$ for all $i \neq j$, $i, j = 1, 2,..., q$. Then, $f(AB) = f(BA)$ implies that there exists an orthogonal matrix Δ such that $\Delta' A \Delta$ and $\Delta' B \Delta$ are diagonal matrices.*

Proof. Since A and B are p.d. matrices, we can find a nonsingular matrix C such that

$$A = CC', \qquad B = C'^{-1} D_b C^{-1}, \qquad AB = CD_b C^{-1}, \tag{2.1}$$

where $D_b = \operatorname{diag}(b_1 I_{p_1}, ..., b_q I_{p_q})$. Then, from the given conditions, it is easy to see that

$$(C'C) f(D_b) = f(D_b)\, C'C, f(D_b) = D_\delta, \tag{2.2}$$

where $D_\delta = \operatorname{diag}(\delta_1 I_{p_1}, ..., \delta_q I_{p_q})$, $\delta_j = f(b_j)$, and $\delta_1 \neq \delta_2 \neq \cdots \neq \delta_q$. From (2.2), it is easy to see that

$$C'C = \operatorname{diag}(C_1, C_2, ..., C_q), \tag{2.3}$$

where C_j ($j = 1, 2,..., q$) are $p_j \times p_j$ p.d. matrices. Hence, it is easy to verify that $C'CD_b = D_b C'C$, i.e., $CD_b C^{-1} = C'^{-1} D_b C'$ or $AB = BA$, which proves the required Lemma 1.

Note 1. If condition (ii) of Lemma 1 is omitted, then the result of Lemma 1 may not be true. For example, consider

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}, \qquad b_1 = 16, \qquad b_2 = 1,$$

and

$$f(x) = (16x^3 - 273x^2 + 273x + 64)\left(\frac{64}{x^4} - \frac{289}{x^3} + \frac{546}{x^2} - \frac{289}{x} + 66\right).$$

Here $f(1) = f(16) = 4000$ and $AB \neq BA$.

LEMMA 2. *Let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$, $u_i = \lambda_i - \lambda_{p-i+1} > 0$ and $v_i = \lambda_i +$*

λ_{p-i+1} *for* $i = 1, 2, \ldots, s$ $(2s < p)$. *Then, for* $s = r + t$ *and for positive integers* $1 \leqslant i_1 < i_2 < \cdots < i_r \leqslant s$,

$$\sum_{i=t+1}^{r+t} (u_i/v_i)^2 \leqslant \sum_{j=1}^{r} (u_{i_j}/v_{i_j})^2 \leqslant \sum_{i=1}^{r} (u_i/v_i)^2,$$

$$\prod_{i=t+1}^{t+r} (u_i/v_i) \leqslant \prod_{j=1}^{r} (u_{i_j}/v_{i_j}) \leqslant \prod_{i=1}^{r} (u_i/v_i),$$

and

$$\sum_{i=t+1}^{t+r} v_i^2/u_i^2 \geqslant \sum_{j=1}^{r} (v_{i_j}/u_{i_j})^2 \geqslant \sum_{i=1}^{r} (v_i/u_i)^2.$$

Proof follows from

$$u_j/v_j \leqslant u_i/v_i \leqslant u_k/v_k \qquad \text{for} \quad k < i < j.$$

LEMMA 3. *Let* A *and* B *be* $p \times p$ p.d. *matrices and let* $X \in \mathscr{M}(p, r)$ *with* $2r \leqslant p$. *Let* $\varphi_1 \geqslant \varphi_2 \geqslant \cdots \geqslant \varphi_p > 0$ *be the eigenvalues of* AB^{-1}. *Then*

$$\prod_{i=p-r+1}^{p} \varphi_i \leqslant | X'AX |/| X'BX | \leqslant \prod_{i=1}^{r} \varphi_i$$

and

$$\sum_{i=p-r+1}^{p} \varphi_i \leqslant \operatorname{tr}\{(X'AX)(X'BX)^{-1}\} \leqslant \sum_{i=1}^{r} \varphi_i \,.$$

Proof. Since A and B are p.d. matrices, we can find a nonsingular matrix C such that

$$B = CC' \qquad \text{and} \qquad A = CD_\varphi C'$$

with $D_\varphi = \operatorname{diag}(\varphi_1, \varphi_2, \ldots, \varphi_p)$. If $Y = C'X\{(X'BX)^{-1/2}\}'$, then

$$Y'Y = I_r, \qquad | X'AX |/| X'BX | = | Y'D_\varphi Y | = \prod_{i=1}^{r} ch_i(Y'D_\varphi Y)$$

and

$$\operatorname{tr}\{(X'AX)(X'BX)^{-1}\} = \operatorname{tr}(Y'D_\varphi Y) = \sum_{i=1}^{r} ch_i(Y'D_\varphi Y)$$

where $ch_1(\cdot) \geqslant ch_2(\cdot) \geqslant \cdots \geqslant ch_r(\cdot)$ are the eigenvalues of $(\cdot)$. The use of Poincare separation theorem (see, for example, Rao (1973, p. 64)), namely,

$$\varphi_{p-i+1} \leqslant ch_i(Y'D_\varphi Y) \leqslant \varphi_i \qquad \text{for} \quad i = 1, 2, \ldots, r$$

gives the required Lemma 3.

LEMMA 4. *Let $X, Y \in \mathscr{M}(p, r, s)$ with $r \leqslant s \leqslant p - r$, and let C and Σ be $r \times r$ and $p \times p$ p.d. matrices. If $B(X, Y) = X'\Sigma Y(Y'\Sigma Y)^{-1}Y'\Sigma X$ and $A(X) = X'\Sigma X - X'X(X'\Sigma^{-1}X)^{-1}X'X$, then*

(i) $\sup_Y |B(X, Y)| = |A(X)|$,

(ii) $\sup_Y \operatorname{tr}(CB(X, Y)) = \operatorname{tr}(CA(X))$,

(iii) $\sup_Y[\operatorname{tr}\{C(B(X, Y))^{-1}\}]^{-1} = [\operatorname{tr}\{C(A(X))^{-1}\}]^{-1}$,

and

(iv) $\sup_Y\{|X'\Sigma X - B(X, Y)|\}^{-1} = |X'\Sigma^{-1}X|/|X'X|^2$.

Proof. Choose $Z \in \mathscr{M}(p, p - r - s)$ such that the columns of $X_1 = \Sigma^{-1/2}X$, $Y_1 = (\Sigma^{1/2})'Y$, and $Z_1 = (\Sigma^{-1/2})Z$ span three mutually orthogonal subspaces whose linear sum is R^p. Then,

$$I = X_1(X_1'X_1)^{-1}X_1' + Y_1(Y_1'Y_1)^{-1}Y_1' + Z_1(Z_1'Z_1)^{-1}Z_1'. \tag{2.4}$$

Let $X_2 = (\Sigma^{1/2})'X$. Then, using (2.4),

$$\begin{aligned} B(X, Y) &= X_2'Y_1(Y_1'Y_1)^{-1}\, Y_1'X_2 \\ &= X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}\, X_1'X_2 - X_2'Z_1(Z_1'Z_1)^{-1}\, Z_1'X_2 \\ &= A(X) - P(X, Z), \end{aligned} \tag{2.5}$$

where $P(X, Z) = X'Z(Z'\Sigma^{-1}Z)^{-1}Z'X$. Note that $A(X)$ is p.d. and $P(X, Z)$ is p.s.d. Hence, it is easy to establish the following:

$$|B(X, Y)| = |A(X) - P(X, Z)| \leqslant |A(X)|,$$

$$\operatorname{tr}(CB(X, Y)) = \operatorname{tr}(CA(X)) - \operatorname{tr}(CP(X, Z)) \leqslant \operatorname{tr}(CA(X)),$$

$$|X_2'X_2 - B(X, Y)| = |(X'\Sigma X - A(X)) - P(X, Z)| \leqslant |(X'\Sigma X - A(X))|,$$

and

$$\operatorname{tr}\{C(B(X, Y))^{-1}\} = \operatorname{tr}\{C(A(X) - P(X, Z))^{-1}\} \geqslant \operatorname{tr}\{C(A(X))^{-1}\}.$$

These results establish the required Lemma 4.

THEOREM 1. *Let* $X, Y \in \mathscr{M}(p, r, s)$ *with* $r \leqslant s \leqslant p - r$ *and let* $B(X, Y) = X'\Sigma Y(Y'\Sigma Y)^{-1}Y'\Sigma X$ *and* $A(X) = X'\Sigma X - X'X(X'\Sigma^{-1}X)^{-1}X'X$. *Then,*

$$\text{(i)} \quad \sup_{X, Y \in \mathscr{M}(p,r,s)} |B(X, Y)|/|X'\Sigma X|$$
$$= \sup_{X \in \mathscr{M}(p,r)} |A(X)|/|X'\Sigma X|$$
$$= \prod_{i=1}^{r} \{(\lambda_i - \lambda_{p-i+1})/(\lambda_i + \lambda_{p-i+1})\}^2,$$

$$\text{(ii)} \quad \sup_{X, Y \in \mathscr{M}(p,r,s)} \operatorname{tr}\{(X'\Sigma X)^{-1} B(X, Y)\}$$
$$= \sup_{X \in \mathscr{M}(p,r)} \operatorname{tr}\{(X'\Sigma X)^{-1} A(X)\}$$
$$= \sum_{i=1}^{r} \{(\lambda_i - \lambda_{p-i+1})/(\lambda_i + \lambda_{p-i+1})\}^2,$$

and

$$\text{(iii)} \quad \sup_{X, Y \in \mathscr{M}(p,r,s)} [\operatorname{tr}\{(X'\Sigma X)(B(X, Y))^{-1}\}]^{-1}$$
$$= \sup_{X \in \mathscr{M}(p,r)} [\operatorname{tr}\{(X'\Sigma X)(A(X))^{-1}\}]^{-1}$$
$$= \left\{\sum_{i=1}^{r} (\lambda_i + \lambda_{p-i+1})^2/(\lambda_i - \lambda_{p-i+1})^2\right\}^{-1}.$$

Proof. The first part of Theorem 1(i) (or Theorem 1(ii) or 1(iii)) follows immediately from Lemma 4(i) (or Lemma 4(ii) or 4(iii)). We have to prove only the second part. For (i), we have

$$g(X) = |A(X)|/|X'\Sigma X|$$

and using the results before Lemma 1, the differential of $\ln g(X)$ can be written as

$$\begin{aligned}(d \ln g(X)) = 2 \operatorname{tr}[&(A(X))^{-1}\{X'\Sigma - X'X(X'\Sigma^{-1}X)^{-1}X'\}(dX) \\ &- (X'\Sigma^{-1}X)^{-1} X'X(A(X))^{-1}\{X' - X'X(X'\Sigma^{-1}X)^{-1} X'\Sigma^{-1}\}(dX) \\ &- (X'\Sigma X)^{-1} X'\Sigma(dX)],\end{aligned}$$

and

$$(d \ln g(X)) = \operatorname{tr}\left(\frac{\partial \ln g(X)}{\partial X}\right)' (dX).$$

Hence, for the stationary values of $g(X)$, we must have

$$[(AX))^{-1} - (X'\Sigma X)^{-1}]\, X'\Sigma + (X'\Sigma^{-1}X)^{-1}X'X(A(X))^{-1}X'X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}$$
$$= [(A(X))^{-1}X'X(X'\Sigma^{-1}X)^{-1} + (X'\Sigma^{-1}X)^{-1}X'X(A(X))^{-1}]\, X'.$$

Substituting

$$X_1 = X\{(X'X)^{-1/2}\}',$$
$$B_2 = \{(X'X)^{1/2}\}'(X'\Sigma^{-1}X)^{-1}(X'X)^{1/2} = (X_1'\Sigma^{-1}X_1)^{-1},$$

and

$$B_1 = (X'X)^{-1/2}\, X'\Sigma X\{(X'X)^{-1/2}\}' = X_1'\Sigma X_1\,,$$

we can write the above equation as

$$(B_1 - B_2)^{-1}\, B_2B_1^{-1}X_1'\Sigma + B_2(B_1 - B_2)^{-1}\, B_2X_1'\Sigma^{-1} = \Lambda X_1'\,, \qquad (2.6)$$

where $X_1'X_1 = I_r$ and $\Lambda = (B_1 - B_2)^{-1}B_2 + B_2(B_1 - B_2)^{-1}$. Premultiplying and postmultiplying (2.6) by $B_1B_2^{-1}(B_1 - B_2)$ and ΣX_1, respectively, we obtain, after some simplifications,

$$X_1'\Sigma^2X_1 = B_1B_2^{-1}B_1B_2 - B_1B_2 + B_1^{\,2}$$

which must be p.s.d. Taking the symmetry into consideration, we get, after minor modifications,

$$(B_1B_2^{-1})^2 - (B_1B_2^{-1}) = (B_2^{-1}B_1)^2 - (B_2^{-1}B_1). \qquad (2.7)$$

For the use of Lemma 1, $f(x) = x^2 - x$. We may observe that the minimum eigenvalue of $B_1B_2^{-1}$ is greater than or equal to one (see, for example, Bloomfield and Watson (1975, p. 122)) and hence the conditions of Lemma 1 are satisfied. This shows that there exists an orthogonal matrix Δ such that $\Delta'B_1\Delta$ and $\Delta'B_2\Delta$ are diagonal matrices. Replacing $X_1\Delta$ by X_2, we see that $X_2'X_2 = I$, $X_2'\Sigma X_2$, and $X_2'\Sigma^{-1}X_2$ are diagonal matrices,

$$g(X) = |\, X_2'\Sigma X_2 - (X_2'\Sigma^{-1}X_2)^{-1}\,|/|\, X_2'\Sigma X_2\,|$$

and (2.6) reduces to

$$(X_2'\Sigma^{-1}X_2)^{-1}\, X_2'\Sigma^{-1} + (X_2'\Sigma X_2)^{-1}\, X_2'\Sigma = 2X_2'$$

which has been solved by Bloomfield and Watson (1975) and Knott (1975). Using their arguments and Lemma 2, we get the required Theorem 1(i).

For the proof of Theorem 1(ii), we proceed in the same manner as above by taking $g(X) = \operatorname{tr}\{(X'\Sigma X)^{-1}(A(X))\}$ and in place of (2.6), we get

$$B_1^{-1}B_2B_1^{-1}X_1'\Sigma + B_2B_1^{-1}B_2X_1'\Sigma^{-1} = (B_1^{-1}B_2 + B_2B_1^{-1})X_1'. \tag{2.8}$$

Then, in place of (2.7), we get

$$(B_1B_2^{-1})^{-3}(B_1B_2^{-1} - I) = (B_2^{-1}B_1 - I)(B_2^{-1}B_1)^{-3}. \tag{2.9}$$

For the use of Lemma 1, $f(x) = x^{-3}(x - 1)$. Let y_1 and y_2 ($<y_1$) be two distinct eigenvalues of $B_2B_1^{-1}$. Then, let us suppose, if possible, $f(1/y_1) = f(1/y_2)$ and this will be possible if

$$y_1^2 + y_2^2 + y_1y_2 = y_1 + y_2. \tag{2.10}$$

The solution of y_2 given y_1 is

$$2y_2 = 1 - y_1 + [(1 + 3y_1)(1 - y_1)]^{1/2}$$

or (2.11)

$$2y_2 = 1 - y_1 - [(1 + 3y_1)(1 - y_1)]^{1/2}.$$

From (2.10), $y_1/y_2 = (y_1 + y_2 - 1)/(1 - y_1) > 1$ or $y_2 > 2(1 - y_1)$. Using this in (2.11), we find that (2.10) will be true if

$$2y_2 = 1 - y_1 + [(1 + 3y_1)(1 - y_1)]^{1/2}.$$

Further, $2y_2$ must be less than $2y_1$. This gives

$$\{(1 + 3y_1)(1 - y_1)\}^{1/2} < 3y_1 - 1$$

or

$$y_1 > 4/3$$

which is impossible because $y_1 \leqslant 1$. Thus, (2.10) does not hold for the permissible values of the roots of $B_2B_1^{-1}$ (or $B_1B_2^{-1}$). Thus, the conditions of Lemma 1 are satisfied and we can find an orthogonal matrix Δ such that $\Delta'B_1\Delta$ and $\Delta'B_2\Delta$ are diagonal matrices. Then arguing as in the case of Theorem 1(i) we get Theorem 1(ii).

For the proof of Theorem 1(iii), we have

$$g(X) = [\operatorname{tr}\{(X'\Sigma X)(A(X))^{-1}\}]^{-1}$$

and in place of (2.6), we get

$$(B_2^{-1}B_1 - I)^{-2}\, B_2^{-1}X_1'\Sigma + (B_1B_2^{-1} - I)^{-2}\, B_1X_1'\Sigma^{-1} = AX_1', \tag{2.12}$$

where

$$A = (B_2^{-1}B_1 - I)^{-2}\, B_2^{-1}B_1 + (B_1B_2^{-1} - I)^{-2}\, B_1B_2^{-1}.$$

Then, in place of (2.7), we get

$$B_1^{-1}B_2(B_2^{-1}B_1 - I)^3 = (B_1B_2^{-1} - I)^3\, B_2B_1^{-1}. \tag{2.13}$$

For the use of Lemma 1, $f(x) = (x - 1)^3/x$. Let y_1 and y_2 ($< y_1$) be two distinct eigenvalues of $(B_2B_1^{-1})$. Then, let us suppose, if possible, $f(1/y_1) = f(1/y_2)$, and this will be possible if

$$(y_2y_1)^2 - 3(y_1y_2) + y_1 + y_2 = 0.$$

The solution of the equation for y_2 is

$$2y_1^2y_2 = 3y_1 - 1 + [(5y_1 - 1)(y_1 - 1)]^{1/2}$$

or

$$2y_1^2y_2 = 3y_1 - 1 - [(5y_1 - 1)(y_1 - 1)]^{1/2},$$

which is imaginary for $\frac{1}{5} < y_1 < 1$ and which is negative for $0 < y_1 \leqslant \frac{1}{5}$. The permissible region of y_1 is (0, 1). Hence, the conditions of Lemma 1 are satisfied and we get $B_1B_2 = B_2B_1$. Then, arguing as in the case of Theorem 1(i) above, we get the required result for Theorem 1(iii). This completes the proof of Theorem 1.

THEOREM 2. *Let $X, Y \in \mathscr{M}(p, r, s)$ and let $Z \in \mathscr{M}(p, 2t)$ be such that $X'Z = 0$ and $Y'Z = 0$ and $r \leqslant s \leqslant p - r$ and $p \geqslant r + s + 2t$. Let $\lambda_{r+t} > \lambda_{p-r-t+1}$, where $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$ are the eigenvalue of Σ. Then,*

$$\text{(i)} \quad \inf_{Z\in\mathscr{M}(p,2t)} \sup_{\substack{X,Y\in\mathscr{M}(p,r,s)\\ X'Z=0 \text{ and } Y'Z=0}} \frac{|B(X, Y)|}{|X'\Sigma X|}$$
$$= \inf_{Z\in\mathscr{M}(p,2t)} \sup_{\substack{X\in\mathscr{M}(p,r)\\ X'Z=0}} \frac{|A(X)|}{|X'\Sigma X|}$$
$$= \prod_{i=t+1}^{r+t} \{(\lambda_i - \lambda_{p-i+1})/(\lambda_i + \lambda_{p-i+1})\}^2;$$

(ii)
$$\inf_{Z\in\mathscr{M}(p,2t)} \sup_{\substack{X,Y\in\mathscr{M}(p,r,s)\\ X'Z=0 \text{ and } Y'Z=0}} \operatorname{tr}\{(X'\Sigma X)^{-1} B(X,Y)\}$$
$$= \inf_{Z\in\mathscr{M}(p,2t)} \sup_{\substack{X\in\mathscr{M}(p,r)\\ X'Z=0}} \operatorname{tr}\{(X'\Sigma X)^{-1} A(X)\}$$
$$= \sum_{i=t+1}^{r+t} \{(\lambda_i - \lambda_{p-i+1})/(\lambda_i + \lambda_{p-i+1})\}^2;$$

and

(iii)
$$\inf_{Z\in\mathscr{M}(p,2t)} \sup_{\substack{X,Y\in\mathscr{M}(p,r,s)\\ X'Z=0 \text{ and } Y'Z=0}} [\operatorname{tr}\{(X'\Sigma X)(B(X,Y))^{-1}\}]^{-1}$$
$$= \inf_{Z\in\mathscr{M}(p,2t)} \sup_{\substack{X\in\mathscr{M}(p,r)\\ X'Z=0}} [\operatorname{tr}\{(X'\Sigma X')(A(X))^{-1}\}]^{-1}$$
$$= \left[\sum_{i=t+1}^{r+t} \{(\lambda_i + \lambda_{p-i+1})/(\lambda_i - \lambda_{p-i+1})\}^2\right]^{-1},$$

where

$B(X, Y) = X'\Sigma Y(Y'\Sigma Y)^{-1}Y'\Sigma X$ *and* $A(X) = X'\Sigma X - X'X(X'\Sigma^{-1}X)^{-1}X'X$.

Proof. The proof of Theorem 2 will be given only for case (i). The results for the other cases can be obtained on the similar lines. We shall establish the result (i) when Σ is a diagonal matrix $D_\lambda = \operatorname{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$, $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p > 0$, because if Σ is not a diagonal matrix, it can be reduced to a diagonal matrix by an orthogonal matrix P, and the problem is unchanged by replacing $P'X$, $P'Y$, and $P'Z$ by X, Y, and Z, respectively. We observe that the infimum over the whole space $Z \in \mathscr{M}(p, 2t)$ is less than or equal to the infimum over the subspace $Z = Z_0$, where

$$Z_0' = \left(\begin{array}{c|c|c} I_t & 0 & 0 \\ \hline 0 & 0 & I_t \end{array}\right).$$

If $X' = (X_1', X_2', X_3')$, $Y' = (Y_1', Y_2', Y_3')$, and $D_\lambda = \operatorname{diag}(D_1, D_2, D_3)$, where $D_2 = \operatorname{diag}(\lambda_{t+1}, ..., \lambda_{p-t})$, then φ, the left-hand side of (i) is given, by

$$\varphi \leqslant \sup_{\substack{X,Y\in\mathscr{M}(p,r,s)\\ X'Z_0=0 \text{ and } Y'Z_0=0}} \frac{|B(X,Y)|}{|X'\Sigma X|}$$
$$= \sup_{X_2,Y_2\in\mathscr{M}(p-2t,r,s)} \frac{|B(X_2,Y_2)|}{|X_2'D_2X_2|}$$
$$= \prod_{i=t+1}^{r+t} \left\{\frac{(\lambda_i - \lambda_{p-i+1})^2}{(\lambda_i + \lambda_{p-i+1})^2}\right\}. \tag{2.14}$$

Now, since the supremum over the whole space $X, Y \in \mathscr{M}(p, r, s)$ is greater than or equal to the supremum over the subspace given by $X = X_0$ and $Y = Y_0$ where

$$X_0 = \begin{array}{c} \begin{pmatrix} X_4 \\ \hline 0 \\ \hline X_5 \end{pmatrix} \\ r \end{array} \begin{array}{l} t+r \\ p-2r-2t, \\ t+r \\ \end{array} \qquad Y_0 = \begin{array}{c} \left(\begin{array}{c|c} X_4 & 0 \\ \hline 0 & Y_1 \\ \hline -X_5 & 0 \end{array}\right) \\ r \quad s-r \end{array} \begin{array}{l} t+r \\ p-2r-2t, \\ t+r \\ \end{array}$$

$$X_4 = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_{t+r}' \end{pmatrix}, \qquad X_5 = \begin{pmatrix} x_{t+r}' \\ x_{t+r-1}' \\ \vdots \\ x_1' \end{pmatrix}$$

with $X_4 \in \mathscr{M}(t + r, r)$ and $Y_1 \in \mathscr{M}(p - 2r - 2t, s - r)$. Note that

$$\frac{| B(X_0 , Y_0)|}{| X_0' D_\lambda X_0 |} = \left(\frac{| X_4' D_u X_4 |}{| X_4' D_v X_4 |}\right)^2,$$

where $u_i = \lambda_i - \lambda_{p-i+1}$, $v_i = \lambda_i + \lambda_{p-i+1}$ for $i = 1, 2,..., t + r$, $D_u = \operatorname{diag}(u_1 ,..., u_{t+r})$, and $D_v = \operatorname{diag}(v_1 , v_2 ,..., v_{t+r})$. Then, using Lemma 3,

$$| B(X_0 , Y_0)|/| X_0' D_\lambda X_0 | \geqslant \prod_{i=t+1}^{t+r} (u_i/v_i)^2 = \prod_{i=t+1}^{t+r} \{(\lambda_i - \lambda_{p-i+1})/(\lambda_i + \lambda_{p-i+1})\}^2.$$

Hence,

$$\varphi \geqslant \inf_{Z \in \mathscr{M}(p,2t)} \sup_{\substack{X_0, Y_0 \text{ with} \\ X_0' Z = 0 \text{ and } Y_0' Z = 0}} \frac{| B(X_0 , Y_0)|}{| X_0' D_\lambda X_0 |}$$

$$\geqslant \prod_{i=t+1}^{t+r} \{(\lambda_i - \lambda_{p-i+1})/(\lambda_i + \lambda_{p-i+1})\}^2. \tag{2.15}$$

From, (2.14) and (2.15), we get

$$\varphi = \prod_{i=t+1}^{t+r} \{(\lambda_i - \lambda_{p-i+1})/(\lambda_i + \lambda_{p-i+1})\}^2. \tag{2.16}$$

The middle equality of (i) follows from Lemma 4. Thus, Theorem 2(i) is established. In the similar ways, the other parts can be established. Thus, Theorem 2 is established.

3. Some Applications

3.1. *Canonical Correlations*

Let the random vector w be partitioned as $w' = (w_1', w_2')$ where w_j is a p_j-vector ($j = 1, 2$) and $p = p_1 + p_2$. Let $V \in \mathscr{M}(p_1, r)$ and $W \in \mathscr{M}(p_2, s)$ with $s \geqslant r$.

If $x_1 = V'w_1$ and $x_2 = W'w_2$, then the three measures g, a, and h for the dependence of x_1 and x_2 are given by

$$g_1(V, W) = | V'\Sigma_{12}W(W'\Sigma_{22}W)^{-1} W'\Sigma_{12}'V |/| V'\Sigma_{11}V |, \tag{3.1}$$

$$a_1(V, W) = \operatorname{tr}\{(V'\Sigma_{11}V)^{-1} V'\Sigma_{12}W(W'\Sigma_{22}W)^{-1} W'\Sigma_{12}'V\}/r, \tag{3.2}$$

and

$$h_1(V, W) = r/\operatorname{tr}[V'\Sigma_{11}V\{V'\Sigma_{12}W(W'\Sigma_{22}W)^{-1} W'\Sigma_{12}'V\}^{-1}], \tag{3.3}$$

where the covariance matrix of w is

$$\Sigma = \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{12}' & \Sigma_{22} \end{array}\right).$$

Now, it is easy to verify that

$$\sup_{V\in\mathscr{M}(p_1,r)} \sup_{W\in\mathscr{M}(p_2,s)} g_1(V, W) = \prod_{i=1}^{r} \theta_i^2, \tag{3.4}$$

$$\sup_{V\in\mathscr{M}(p_1,r)} \sup_{W\in\mathscr{M}(p_2,s)} a_1(V, W) = \sum_{i=1}^{r} \theta_i^2/r, \tag{3.5}$$

and

$$\sup_{V\in\mathscr{M}(p_1,r)} \sup_{W\in\mathscr{M}(p_2,s)} h_1(V, W) = r\Big/\sum_{i=1}^{r} \theta_i^{-2} \tag{3.6}$$

where $\theta_1^2 \geqslant \theta_2^2 \geqslant \cdots \geqslant \theta_p > 0$ are the square of the canonical correlations between w_1 and w_2.

Taking $V_*' = (V', 0)$ and $W_*' = (0, W')$ in (3.1)–(3.3), we find that

$$g_1(V, W) = g(V_*, W_*),\ \ a_1(V, W) = a(V_*, W_*),\ \text{ and }\ h_1(V, W) = h(V_*, W_*), \tag{3.7}$$

where g, a, and h are defined by (1.8), (1.9), and (1.10), respectively. Now, using (3.4)–(3.6) and Theorem 1, we get the following inequalities

$$\prod_{i=1}^{r} \theta_i^2 \leqslant \prod_{i=1}^{r} \{(\lambda_i - \lambda_{p-i+1})^2/(\lambda_i + \lambda_{p-i+1})^2\}, \tag{3.8}$$

$$\sum_{i=1}^{r} \theta_i^2 \leqslant \sum_{i=1}^{r} \{(\lambda_i - \lambda_{p-i+1})^2/(\lambda_i + \lambda_{p-i+1})^2\}, \tag{3.9}$$

and

$$\sum_{i=1}^{r} \theta_i^{-2} \geqslant \sum_{i=1}^{r} \{(\lambda_i + \lambda_{p-i+1})^2/(\lambda_i - \lambda_{p-i+1})^2\}. \tag{3.10}$$

When $r = 1$, we get the result established by Eaton (1976). As a consequence of (3.8) and (3.9) for $r = 2$, we get

$$\theta_1 + \theta_2 \leqslant (\lambda_1 - \lambda_p)(\lambda_1 + \lambda_p)^{-1} + (\lambda_2 - \lambda_{p-1})(\lambda_2 + \lambda_{p-1})^{-1}$$

and using (3.8) and (3.10) for $r = 2$, we get

$$\left\{\frac{\lambda_1 - \lambda_p}{\lambda_1 + \lambda_p} + \frac{\lambda_2 - \lambda_{p-1}}{\lambda_2 + \lambda_{p-1}}\right\} \theta_1\theta_2 \leqslant (\theta_1 + \theta_2)\left(\frac{\lambda_1 - \lambda_p}{\lambda_1 + \lambda_p}\right)\left(\frac{\lambda_2 - \lambda_{p-1}}{\lambda_2 + \lambda_{p-1}}\right).$$

3.2. *Union–Intersection Test Procedures for Testing the Sphericity*

Let x be $N(\mu, \Sigma)$ and let there be n independent observations on x. It has been shown by Mallows (1961) that the hypothesis $H(\Sigma = \sigma^2 I)$ against $H(\Sigma \neq \sigma^2 I)$ is equivalent to testing the independence of two vectors $X'x$ and $Y'x$ where $X \in \mathcal{M}(p, r)$, $Y \in \mathcal{M}(p, s)$, and $X'Y = 0$ for every $s \geqslant r = 1, 2, \ldots, p/2$. The sample correlation matrix R between $X'x$ and $Y'x$ is given by

$$R = (X'SX)^{-1/2}(X'SY)\{(Y'SY)^{-1/2}\}',$$

where S is the sample covariance matrix and $(n - 1)S$ is distributed as Wishart under H_0. We propose to reject $H_0(\Sigma = \sigma^2 I)$ for large values of $|RR'|$ or $\mathrm{tr}(RR')/r$ or $r/\mathrm{tr}(RR')^{-1}$. If $l_1 \geqslant l_2 \geqslant \cdots \geqslant l_p > 0$ are the eigenvalues of S, then using Theorem 1, the union–intersection test statistics are given by

$$t_1 = \prod_{i=1}^{r} \{(l_i - l_{p-i+1})^2/(l_i + l_{p-i+1})^2\}, \tag{3.11}$$

$$t_2 = \sum_{i=1}^{r} \{(l_i - l_{p-i+1})^2/(l_i + l_{p-i+1})^2\}/r, \tag{3.12}$$

and

$$t_3 = r\left[\sum_{i=1}^{r}\{(l_i + l_{p-i+1})^2/(l_i - l_{p-i+1})^2\}\right]^{-1}. \tag{3.13}$$

Note that the above three statistics are different from those given by Venables (1976) and this is due to the reason that he maximized $v = |I - RR'|^{-1}$, the "likelihood ratio" measure.

Theorem 2 can be utilized when we are interested in testing the intermediate roots, but we shall not express this explicitly.

Acknowledgments

The author is thankful to the referees for some useful suggestions, and to one of them for pointing out an error in the old version of Lemma 1.

References

[1] Amir-Moez and Ali, R. (1956). Extreme properties of eigen values of Hermitian transformation and singular values of the sum and product of linear transformations. *J. Duke Math.* **23** 463–476.

[2] Bloomfield, P., and Watson, G. S. (1975). The inefficiency of least squares. *Biometrika* **62** 121–128.

[3] Eaton, M., L. (1976). A maximization problem and its application to Canonical Correlation. *J. Multivariate Anal.* **6** 422–425.

[4] Knott, M. (1975). On the minimum efficiency of least squares. *Biometrika* **62** 129–132.

[5] Rao, C. R. (1973). *Introduction to Linear Statistical Inference.* 2nd ed. Wiley, New York.

[6] Venables, W. (1976). Some implications of the union intersection principle for tests of sphericity. *J. Multivariate Anal.* **6** 185–190.